UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/719,820	11/21/2003	Feng-Wei Chen Russell	RSW920030184US1	3330

23550 7590 05/03/2006

HOFFMAN WARNICK & D'ALESSANDRO, LLC
75 STATE STREET
14TH FL
ALBANY, NY 12207

EXAMINER	
ROSE, HELENE ROBERTA	
ART UNIT	PAPER NUMBER
2163	

DATE MAILED: 05/03/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/719,820	RUSSELL ET AL.	
	Examiner	Art Unit	
	Helene Rose	2163	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *November 21, 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-22* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-22* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *21 November 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *November 21, 2003*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

Detailed Action

1. Claims 1-22 have been presented for examination.

2. Claims 1-22 have been rejected.

Information Disclosure Statement

3. The information disclosure statement filed 11/21/2003 is objected because "**JP 10065159**" and "**903873**" does not indicate on the Information Disclosure Statement form "what is to be considered", whether it's the abstract or full text document. In this case, the Examiner has objected to these two foreign documents because it fails to comply with 37 CFR 1.98(a)(2), which requires a legible copy of each cited foreign patent document; each non-patent literature publication or that portion which caused it to be listed; and all other information or that portion which caused it to be listed.

In reference to the US Patents and the Non-Patent Literature cited on the Information Disclosure Statement, Examiner is considering documents as well as the Foreign Patent document "GB 2227106A" because it provides a full translation of the abstract and full text document. It has been placed in the application file, but the information referred to therein has not been considered.

Claim Objections

4. Claims 11 and 20 are objected to because of the following informalities: Claims 11 and 20 have parentheses within claims, Examiner suggest removing the cited parentheses from claim limitation. Appropriate correction is required.

Claim Rejections – 35 U.S.C – 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 1-22 are rejected under 35 U.S.C. 102(e) as being anticipated by Miller et al (US Patent No. 6,687,695, Filing Date of Patent: October 1, 1999).

Claims 1 and 19:

Regarding claims 1 and 19, Miller teaches a program product stored on a recordable medium for generating a data mining model (Figure 1, all features and wherein it further defined in column 4, lines 33-67, Miller), which when executed comprises:

program code for generating a plurality of datasets from sample data (columns 6-7, lines 55-67 and lines 1-16, wherein data description is defined to be the generation of descriptive statistics, frequencies, and/or histogram bin, and wherein data derivation is defined to be generate new variables such as predictive variables such as bitmaps, ranges, codes, and mathematical functions, and wherein data reduction is defined to be the ability to reduce the number of variables or observations used when designing an analytic model, and so forth, and wherein a sequence of data mining operations is viewed as set of steps that start with some collection of table in a database, generate series of intermediate work tables, and finally produce a result table or view, wherein a data set is interpreted to

be all data belonging to a particular observation block including UV data, astronomical information, i.e. sources, coordinates, frequencies, etc., reduction scripts, system data, logs, warnings, and environmental data, etc. and columns 7-8, lines 56-67, wherein set of tables in the relational database that are populated with transaction level data the source of which could point of sale devices, automated teller machines, call center, and so forth is interpreted to be the sample data, Miller);

program code for selecting a set of algorithms based on objectives for the data mining model (column 4, line 61, wherein selecting the data set and "pre-processing" the data, Miller);

program code for optimizing the set of algorithms using the plurality of datasets (column 5, lines 20-22, wherein one or more scalable data mining functions are optimized SQL statements that perform advanced analytic processing in the RDBS, wherein the analytic application program provides a mechanism for an APPL or component to invoke the scalable data mining functions, Miller); and

program code for generating the data mining model based on the optimized set of algorithms (column 5, lines 61-67, wherein client interacts with the analytic algorithm application program, i.e. APPL which interfaces to the analytic API to invoke one or more of the scalable data mining functions which are executed by the RDBMS, wherein the results from the execution of the scalable data mining functions would be stored as an analytic model within an analytic LDM in the RDBMS, column 6, lines 44-48, wherein the scalable data mining functions comprise complex optimized SQL statements that are created by parameterizing and instantiating the corresponding analytic API's, lines 52-57 and column 7, lines 1-12, wherein the scalable data mining function can be categorized

by the following functions are defined and a sequence of data mining operations is viewed as set of steps that start with some collection of table in a database, generate series of intermediate work tables, and finally produce a result table or view – interpreted to be the optimized set and column 8, lines 18-39, wherein testing and application and where the rules are validated and applied against new data set, and wherein first extract the data from the database to construct a flat file and then execute the "train" portion on this resultant file and wherein computations, aggregations and or ordering can be run parallel because the nature of the RDBMS, Miller).

Claims 2 and 22:

Regarding claims 2 and 22, Miller teaches wherein the program code for generating the plurality of datasets includes:

program code for shuffling the sample data (column 8, lines 11-14, wherein analytic algorithms that require a mix of programmatic iteration along with Extended ANSI SQL statements – is interpreted and equivalent to shuffling, Miller);

program code for placing the shuffled sample data into a plurality of partitions (column 7, lines 3-5, wherein data sampling/partitioning, wherein the ability to intelligently request different data samples or data partitions, Miller); and

program code for including each partition in one of the plurality of datasets (column 7, lines 6-16, wherein scalable data mining functions is to facilitate analytic operations within the RDBMS which process data collections stored in the database and produce results that also are stored in the database, the database comprises a combined storage and workspace environment, as such a sequence of data mining operations is

viewed as a set of steps that start with some collections of tables in the database, generate series of intermediate work tables, and finally produce a result table or view, Miller).

Claim 3:

Regarding claim 3, Miller teaches wherein the plurality of datasets includes a training dataset, a validation dataset, and a testing dataset (column 8, lines 18-39, wherein testing and application and where the rules are validated and applied against new data set, and wherein first extract the data from the database to construct a flat file and then execute the "train" portion on this resultant file and wherein computations, aggregations and or ordering can be run parallel because the nature of the RDBMS, Miller).

Claim 4:

Regarding claim 4, Miller teaches wherein the creating step further includes repeating the including step until each partition is included in at least one training dataset (column 7, lines 22-32, wherein program iteration is defined and interpreted to be repeatedly performing the same sequence of steps and wherein pre-processing of data that reduces the amount of data that a non-SQL algorithm can then process, Miller).

Claim 5:

Regarding claim 5, Miller teaches wherein the selecting step includes obtaining a rule that comprises a best practice for an objective (columns 9-10, line 21, wherein applied against new data set and lines 42-45, wherein scaled data set can be regression or clustering, wherein clustering is interpreted to be a bunch, Miller).

Claim 6:

Regarding claim 6, Miller teaches wherein the best practice is based on at least one of: research, data characteristics, and user feedback (column 8, lines 66-67, wherein

define the characteristics of data stored in the relational database – interpreted to be the data characteristics, Miller).

Claim 7:

Regarding claim 7, Miller teaches wherein the selecting step includes analyzing an attribute of the sample data, and wherein the set of algorithms is further selected based on the attribute (column 8, lines 60-67, wherein the RDBMS provides logical entity and attribute definitions for advanced analytic processing, i.e. scalable data mining functions and analytic algorithms performed by the RDBMS, Miller).

Claim 8:

Regarding claim 8, Miller teaches wherein the optimizing step includes:

applying the set of algorithms to the plurality of datasets (column 7, lines 41-42, wherein analytic algorithms provide data analyst with an unprecedented option to train and apply machine learning analyst against massive amounts of data in the relational database, Miller); and

analyzing a set of results for the applying step (column 8. lines, wherein learns various rules based upon data description followed by testing and application and where the rules are validated and applied to a new data set, wherein this class of algorithms are compute-intensive and historically can not handle large volumes of data because they expect the analyzed data to be in a specific fixed or variable flat file format, Miller).

Claim 9:

Regarding claim 9, Miller teaches wherein the optimizing step further includes:

adjusting at least one algorithm based on the set of results (column 9, lines 7-11, Miller); and

applying the adjusted set of algorithms to the plurality of datasets (columns 9-10, line 21, wherein applied against new data set and lines 42-45, wherein scaled data set can be regression or clustering, wherein clustering is interpreted to be a bunch, Miller).

Claims 10 and 21:

Regarding claims 10 and 21, Miller teaches wherein the program code for generating the data mining model includes program code for generating a set of standard query language statements based on the optimized set of algorithms (column 4, lines 1-14, wherein SQL statements are defined and what statements are utilized, Miller).

Claims 11 and 20:

Regarding claims 11 and 20, Miller teaches program code for storing the data mining model as a character large object (CLOB) in a database (column 4, lines 36-37, wherein massive parallel processing is defined to be a computer hardware architecture designed to obtain high performance through the use of large numbers, i.e. tens, hundreds or thousands, of individually simple, low powered processors, each with its own memory, which is equivalent to CLOB, which is defined to be a term for a large file of characters stored as part of a database record, Miller).

Claim 12:

Regarding claim 12, Miller teaches a computerized method of generating a data mining model, the method comprising:

obtaining a set of algorithms and a plurality of datasets;

applying the set of algorithms to the plurality of datasets;

analyzing a set of results for the applying step;

adjusting at least one algorithm based on the set of results;

applying the adjusted set of algorithms to the plurality of datasets; and

generating the data mining model based on the adjusted set of algorithms.

Examiner Note: Claim 12 have substantially the same limitations as claims 1, 8 and 9. Theses limitations have already been addressed in the rejection of claims 1, 8 and 9. Therefore, they are rejected on the similar grounds corresponding to the discussion above to the rejected claims 1, 8 and 9 above.

Claim 13:

Regarding claim 13, Miller teaches wherein the obtaining step includes:

obtaining sample data (Figure 5, diagram 500 and column 9, lines 66-67, wherein receiving a query or other SQL statements Miller); and

automatically generating the plurality of datasets from the sample data (column 6, lines 3-5, wherein the analytic algorithms comprise SQL statements that **may** or may not include program iteration and the SQL statements are executed by the RDBMS and column 10, lines 3-5, wherein generating a plan that enables the RDBMS to retrieve the correct information from the relational database to satisfy the query, Miller).

Claim 14:

Regarding claim 14, Miller teaches wherein the obtaining step includes:

obtaining objectives for the data mining model (column 7, lines 58-67, wherein relationships it describes relationship between data elements in table and interpreted to be the objectives and column 9, lines 66-67, wherein receiving a query or other SQL statements, Miller); and

automatically selecting the set of algorithms based on the objectives (column 8, lines 1-9, wherein dimension of time can be added in such a way that these relationships

can be analyzed to determine how they change over time, and wherein the implementation is solely within the SQL statements, Miller).

Claim 15:

Regarding claim 15, Miller teaches a system for generating a data mining model, the system comprising:

a dataset system for obtaining a plurality of datasets (**REFER** to claim 1, wherein this limitation has been addressed, Miller);

a rules system for obtaining a plurality of algorithms (column 8, lines 18-21, wherein data mining application, i.e. APPL require a training phase where the APL learns various rules based upon data description, Miller);

an optimization system for optimizing the set of algorithms using the plurality of datasets (**REFER** to claim 1, wherein this limitation has been addressed, Miller); and

a model system for generating the data mining model based on the optimized set of algorithms (column 3, lines 2-6; column 5, lines 20-51, wherein optimization of algorithms is defined, Miller).

Claim 16:

Regarding claim 16, Miller teaches a storage system for storing the data mining model in a database (column 4, lines 46-59, Miller).

Claim 17:

Regarding claim 17, Miller teaches wherein the dataset system automatically generates the plurality of datasets from sample data (column 5, lines 41-45, wherein program iteration is defined and wherein providing a mechanism for an APPL or other

components to invoke the analytic algorithms, Miller)

Claim 18:

Regarding claim 18, Miller teaches wherein the rules system automatically selects the set of algorithms based on objectives for the data mining model (column 3, lines 55-67, wherein data mining process is an interactive approach referred to by Knowledge Discovery Analytic Process, i.e. KDAP, wherein the task our defined, and column 4, lines 1-30, wherein the components to these task are defined, Miller).

Prior Art of Record

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

1. Miller et al	(US Patent No. 6,687,695 and US Patent No. 6,553,366)
2. Tate	(US Patent No. 6,704,717 and US Patent No. 6,611,829)
3. Vishnubhotla	(US Patent No. 6,799,181)

Point of Contact

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Helene Rose whose telephone number is (571) 272-0749. The examiner can normally be reached on 8:00am - 4:30pm Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Don Wong can be reached on (571) 272-1834. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Helene Rose
Technology Center 2100
April 28, 2006

ALFORD KINDRED
PRIMARY EXAMINER